

PATENT ABSTRACTS OF JAPAN

(11)Publication number : 08-227406

(43)Date of publication of application : 03.09.1996

(51)Int.Cl.

G06F 15/16

G06F 15/16

G06F 11/16

G06F 13/00

(21)Application number : 07-033329

(71)Applicant : HITACHI LTD

(22)Date of filing : 22.02.1995

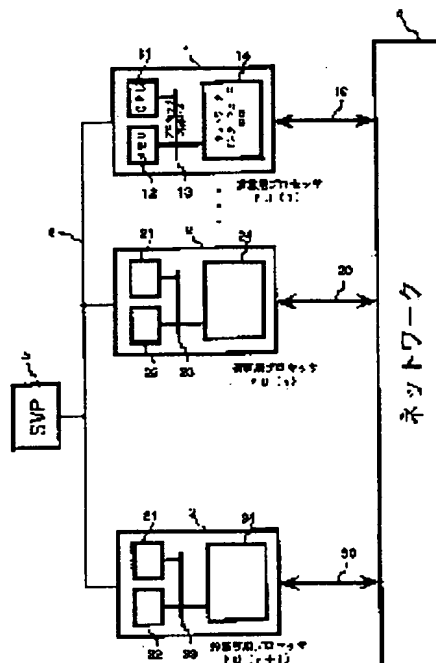
(72)Inventor : AKESE TAKAHISA
ANDO TOSHIMITSU

(54) PARALLEL COMPUTER

(57)Abstract:

PURPOSE: To improve the reliability of a network over which plural processors are connected and perform inter-processor data transfer.

CONSTITUTION: The parallel computer consists of arithmetic processors PU(1) 1... PU(n) 2 and a processor PU(n+1) 3 exclusively for diagnosis which include CPUs 11, 21, and 31, memories 12, 22, and 32, network interface circuits 14, 24, and 34 each having a mechanism that uses a diagnostic packet for diagnosing a fault of a process or network and reporting it to an OS and a user and takes a self-test wherein the diagnostic packet is sent and received by itself, a mechanism that sends and receives the diagnostic packet when the parallel computer system is started up, and a mechanism which sends and receives the diagnostic packet during system operation when the network is not used and periodically and the connected to one another by internal buses 13, 23, and 33, data transfer signals 10, 20, and 30, a service processor(SVP) 5 which is used for the operation of the whole system, and a processor control signal 6.



LEGAL STATUS

[Date of request for examination]

[Date of sending the examiner's decision of rejection]

[Kind of final disposal of application other than the examiner's decision of rejection or application converted registration]

[Date of final disposal for application]

[Patent number]

[Date of registration]

[Number of appeal against examiner's decision of rejection]

[Date of requesting appeal against examiner's decision of rejection]

[Date of extinction of right]

Copyright (C); 1998,2000 Japanese Patent Office

(19) 日本国特許庁 (J P)

(12) 公開特許公報 (A)

(11) 特許出願公開番号

特開平8-227406

(43) 公開日 平成8年(1996)9月3日

(51) Int.Cl. ⁶	識別記号	庁内整理番号	F I	技術表示箇所
G 0 6 F 15/16	4 5 0		G 0 6 F 15/16	4 5 0 Z
	3 7 0			3 7 0 N
11/16	3 1 0		11/16	3 1 0 A
13/00	3 5 1	7368-5E	13/00	3 5 1 N

審査請求 未請求 請求項の数 5 O L (全 13 頁)

(21) 出願番号 特願平7-33329

(22) 出願日 平成7年(1995)2月22日

(71) 出願人 000005108

株式会社日立製作所

東京都千代田区神田駿河台四丁目6番地

(72) 発明者 明瀬 貴久

神奈川県秦野市堀山下1番地 株式会社日

立製作所汎用コンピュータ事業部内

(72) 発明者 安藤 利光

神奈川県秦野市堀山下1番地 株式会社日

立製作所汎用コンピュータ事業部内

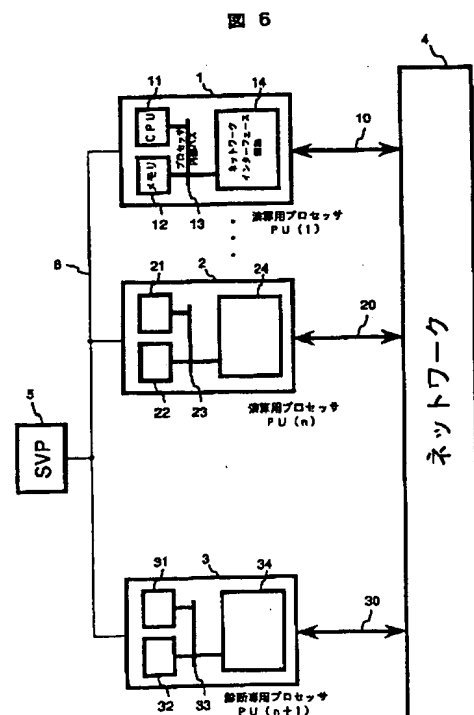
(74) 代理人 弁理士 小川 勝男

(54) 【発明の名称】 並列計算機

(57) 【要約】

【目的】並列計算機において、複数のプロセッサを接続しプロセッサ間データ転送を行うネットワークの信頼性を向上する。

【構成】CPU 11, 21, 31、メモリ 12, 22, 32 及びプロセッサやネットワークの故障を診断しOS やユーザーに対して報告を行う診断パケットを使用し、診断パケットを自分に対して送受信するセルフテストを行う機構、並列計算機システムの立ち上げ時に診断パケットの送受信を行う機構、システム稼働中にネットワークが未使用時及び定期的に診断パケットを送受信する機構を持ったネットワークインターフェース回路 14, 24, 34 を内蔵し、それぞれを内部バス 13, 23, 33 によって接続した演算プロセッサPU (1) 1...PU (n) 2 及び診断専用プロセッサPU (n+1) 3 とプロセッサ間データ転送用ネットワーク 4 とデータ転送信号 10, 20, 30 とシステム全体の運用に使用するサービスプロセッサ (SVP) 5 とプロセッサ制御信号 6 によって構成する。



【特許請求の範囲】

【請求項1】複数のプロセッサを接続したネットワークによりプロセッサ間データ転送を行う並列計算機において、実際の計算に用いるプロセッサ間データ通信とは異なる、ネットワーク上のハードウェアが故障しているか否かを検出する為の診断用パケットをネットワーク上に転送して各プロセッサとネットワーク間の診断を行い障害部位の指摘とOSやユーザーに対する報告を行う並列計算機。

【請求項2】複数のプロセッサを接続したネットワークによりプロセッサ間データ転送を行う並列計算機において、各プロセッサが診断用パケットを自分に対して転送し、セルフテストを行う請求項1に記載されている並列計算機。

【請求項3】複数のプロセッサを接続したネットワークによりプロセッサ間データ転送を行う並列計算機において、実際の計算に用いる演算用プロセッサ以外にネットワーク上に診断専用のプロセッサを別に接続し、このプロセッサがネットワーク又は他のプロセッサに対して診断用パケットの送受信を行い、ネットワークを構成するハードウェアや他のプロセッサの診断を行う請求項1に記載されている並列計算機。

【請求項4】複数のプロセッサを接続したネットワークによりプロセッサ間データ転送を行う並列計算機において、システム立ち上げ時にプロセッサがネットワーク上に診断パケットを転送しネットワークを構成するハードウェアや他のプロセッサの診断を行う請求項1に記載されている並列計算機。

【請求項5】複数のプロセッサを接続したネットワークによりプロセッサ間データ転送を行う並列計算機において、システム運行中ネットワーク未使用時又はタイマー等の使用により定期的に診断パケットをネットワーク上に転送し、ネットワークを構成するハードウェアや他のプロセッサの診断を行う請求項1に記載されている並列計算機。

【発明の詳細な説明】

【0001】

【産業上の利用分野】本発明は、複数のプロセッサを接続したネットワークによりプロセッサ間データ転送を行う並列計算機のネットワーク診断方法に関連する。

【0002】

【従来の技術】従来の複数のコンピュータをLAN等のコンピュータネットワークで接続したシステムにおいて、例えばイーサネットに接続した場合ネットワークに障害があるか否かをテストする従来技術として、W. リチャード・ステーグンス著「UNIXネットワークプログラミング」（トッパン刊）のP527～550にUNIXシステムにおけるpingルーチンが記述されている。

【0003】

【発明が解決しようとする課題】従来技術であるUNIXシステムにおけるpingコマンドは、複数のプロセッサをネットワークで接続しプロセッサ間データ転送を行う並列計算機については考えられていない。また、pingコマンドは、ユーザがマニュアルでコマンドを発行することにより有効になるもので、システム運行中、不注意によりプロセッサ間データ信号線を切断してしまった場合や、何らかの外乱により何れかのプロセッサやネットワークを構成するハードウェアが故障してしまった場合などは障害部位の発見が難しい。プロセッサの数が少なければ従来の方法やサービスプロセッサ等の使用により人手で障害部位の指摘は可能だが、超並列計算機の様にプロセッサの数が多くなってくると障害部位の発見にかなり時間と工数がかかってしまう。

【0004】多くの並列計算機は複数のプロセッサ間を大量のケーブルで接続している可能性が高く、プロセッサ数の増加に比例してケーブルも増えるので、並列計算機は今までの計算機に比べ処理能力が大幅に向上する反面、障害処理を強化しなければ、システムの稼働率が落ちてしまう可能性がある。

【0005】

【課題を解決するための手段】上記課題を解決する為に、複数のプロセッサを接続したネットワークによりプロセッサ間データ転送を行う並列計算機において、従来のネットワーク診断テスト手段を各プロセッサが自分に対して自動的に行うセルフテストの制御手段を持つ。

【0006】またネットワーク上に通常の演算を行うプロセッサ以外に診断専用のプロセッサを設け、この診断専用プロセッサが従来のネットワーク診断テスト手段を自動的に行う制御手段を持つ。

【0007】また、従来のネットワーク診断テスト手段を、システム立ち上げ時に各プロセッサが自動的に行う制御手段を持つ。

【0008】また、従来のネットワーク診断テスト手段を、システム運行中ネットワーク未使用時又は定期的に自動的に行う制御手段を持つ。

【0009】

【作用】本発明に係わる並列計算機において、プロセッサ間ネットワークとネットワークに接続される複数のプロセッサとの間において、ネットワークの故障を診断しOSやユーザーに対して報告を行う診断パケットを使用し、各プロセッサが自分に対して自動的に診断テストを行いセルフテストをする事により、自分がネットワークへのデータ送受信処理を正常に行えるか自覚する事が可能となる。

【0010】また、ネットワーク上に通常の演算を行うプロセッサ以外に診断専用のプロセッサを設け、このプロセッサにのみ自動的に診断を行わせる事により、他の演算プロセッサの演算処理を妨げることなくネットワークの診断が行え、診断テストによるシステム全体の処理

能力の低下を防ぐ事が可能となる。

【0011】また、システム立ち上げ時に各プロセッサが自動的に診断テストを行う事により、何れかのプロセッサ又はネットワークを構成するハードウェアが故障したままシステムが動作を開始してしまい誤動作する事を未然に防ぐ事が可能となる。

【0012】また、システム運行中に各プロセッサがネットワーク未使用時又は定期的に診断テストを自動的に行う事により、システムの運行中にユーザーがコマンドを発行することなくネットワークの診断が行え、何らかの外乱により何れかのプロセッサやネットワークを構成するハードウェア等が故障しても早期に報告されるので、故障したままシステムを運用し続け、誤動作してしまう事を未然に防ぐ事が可能となる。

【0013】

【実施例】以下、本発明の実施例を図を用いて説明する。

【0014】先ず、本発明に係る並列計算機の構成について述べる。

【0015】図1は、本発明に係る並列計算機であり、 n 個の演算プロセッサをネットワークに接続し、プロセッサ間データ転送を行い、かつ、実際の計算に用いるプロセッサ間データ通信とは異なる、ネットワーク上のハードウェアが故障しているか否かを検出する為のネットワーク診断専用通信も行う並列計算機の構成を示す。図1において、1, 2はプロセッサ(PU(1), ..., PU(n))を示す。各プロセッサ1, 2はCPU11, 21とメモリ12, 22及びネットワークインターフェース回路14, 24を備え、それぞれがプロセッサ内部バス13, 23を介して接続される。各プロセッサは、データ転送信号10, 20を介してプロセッサ間結合のネットワーク4に接続されている。5はサービスプロセッサ(SVP)であり、各プロセッサ1, 2に対してプロセッサ制御信号6を介して接続される。

【0016】図2は図1で述べたプロセッサを構成する回路の一つであり、各プロセッサ1, 2とプロセッサ間結合ネットワーク4との間でデータ転送を行う為のインターフェースであるネットワークインターフェース回路14の内部構成を示す。図2において、ネットワークインターフェース回路14は、プロセッサ内メモリの送信パケット領域からデータを読み取り、ネットワークヘッダを送信したり、診断パケットを自動生成する機能を持つ送信回路101と、101とは逆にネットワークからデータを受信しプロセッサ内メモリの受信パケット領域へデータを書き込んだり診断パケットを受信した時、診断回路へ診断パケットを転送する機能を持つ受信回路102と、ネットワークインターフェース回路を制御する為の制御コマンドを保持する為のネットワークコントロールレジスタ(NCR)104と、ネットワークインターフェース回路内部の状態を示すネットワークステータスレジスタ(NSR)103と、受信した診断パケットを解析し、ネットワークを診断するネットワーク診断回路106と、他プロセッサから診断パケットを受信した場合に、パケットヘッダーを書き換え、ネットワークを介して元のプロセッサへ送り返す為のパケットヘッダー書き換え回路107及び時間を監視するタイマー監視回路105を備える。

【0017】送信回路101は、バスデータ入力信号110を介してプロセッサ内部メモリ上のパケット送信領域からパケットを受け取ると、パケット送信信号111にてネットワークへと送信する。また、バスデータ入力信号110を介して内部バス13からOS等が作成した診断パケットが入力されたり、送信回路内部で診断パケットを生成した場合、送信診断パケット出力信号112を介して診断回路106へ出力し、送信診断パケットがパケットヘッダー変換回路107から113の送信診断パケット入力信号を介して入力されると、無条件にパケットヘッダーを書き換えただけの折り返しパケットをパケット送信信号111を介してネットワークへ出力する機能を持つ。さらに、定期的に診断を行う設定にするとタイマー監視回路105からのタイマー信号により、一定の時間毎に診断パケットの送出を行う機能も持つ。また、送信診断パケットを一度送出した後ネットワークステータスレジスタ(NSR)の値により受信診断パケットを受け取るか、受信診断パケットが消失してタイムアウトになる等、診断処理が終了するまで次の診断パケットは出力しない。

【0018】受信回路102は、ネットワークよりパケット受信信号121を介してパケットを受信すると、バスデータ出力信号120を介してプロセッサ内部バスを経由してメモリ上のパケット受信領域へパケットを出力する。また、パケット受信信号121を介して他プロセッサからの診断パケットや自プロセッサの送信回路101が送信した診断パケットを受信すると、他のプロセッサからの診断パケットの場合は折り返しパケット出力信号123を介しパケットヘッダー変換回路107へパケットを出力し、自プロセッサ内部の送信回路101が送信した診断パケットの場合は受信診断パケット出力信号122を介して診断回路106へパケットを出力する機能を持つ。

【0019】ネットワークステータスレジスタ(NSR)103は、送信回路101からの送信回路ステータス信号141や受信回路102からの受信回路ステータス信号142及び診断回路106からの診断回路ステータス信号143のデータを集約し、プロセッサ内部バス13と送信回路101へネットワークステータス信号140を出力する。

【0020】ネットワークコントロールレジスタ(NCR)104は、プロセッサ内部バス13からネットワークコントロール信号130を介してネットワークコントロール

ロール命令を受け、送信回路制御信号131を送信回路101へ出力する。

【0021】診断回路106は、送信回路101からの送信診断バケットと受信回路102からの受信診断バケットを比較し判定を行いネットワークステータスレジスタ(NSR)103へ結果を報告する。この時、タイマー監視回路105からのタイマー信号150を使用して、送信診断バケットが入力されてから一定期間内に受信診断バケットが入力されなければタイムアウトエラーとしてネットワークステータスレジスタ(NSR)103に結果報告する。報告する内容は、診断を行ったプロセッサアドレスと診断結果であり、正常終了かバケットコンペアエラーかタイムアウトエラーである。

【0022】タイマー監視回路105は、一定期間毎にプロセッサ内部の回路と送信回路101、診断回路106へタイマー信号150を出力する。タイマー割り込み発生回路としての機能がある。

【0023】バケットヘッダー書換回路107は受信回路102から折り返しバケット出力信号123を介して診断バケットが入力されると、バケットの種類を送信診断バケットから受信診断バケットに変換し、同時に送信元プロセッサアドレスと受信先プロセッサアドレスをスワップして、折り返しバケット入力信号113を介して送信回路101にバケットヘッダー書き換え後の診断バケットを出力する機能を持つ。

【0024】図3は図2で述べたネットワークインターフェース回路14を構成する回路の一つである送信回路101の内部構成を示す。図3において、送信回路101は、送信回路制御信号131、ネットワークステータス信号140及びタイマー信号により送信回路101内部を制御する送信回路制御回路1101、バケットをネットワークへ出力する送信バッファ1102、診断バケットを送信回路101内部で自動的に作成する診断バケット作成回路1103及びOSなどの作成したメモリ上のバケット内のフラグを読み取り、そのバケットが通常バケットか診断バケットかを判定するバケットコード解析回路1104を備える。

【0025】又、それらの大規模な回路の他に、プロセッサ内部バス13からのバケットが通常バケットか診断バケットかにより送信経路を切り換える通常／診断切り換えスイッチ1111、診断バケットの作成方法を切り換える診断バケット作成手段切り換えセクタ1112、通常バケットを出力するか診断バケットを出力するか切り換える通常／診断バケットセクタ1113、折り返しバケットと通常／診断バケットのどちらを出力するか切り換える折り返しバケットセクタ1114も備える。

【0026】送信回路制御回路1101は、ネットワークコントロールレジスタ(NCR)104からの送信回路制御信号131とネットワークステータスレジスタ

(NSR)103からのネットワークステータス信号140とタイマー監視回路105からのタイマー信号150が入力され内容を解析し、自動診断バケット作成信号1161を自動診断バケット作成回路1104に、診断バケット作成手段切り換え信号1162を診断バケット作成手段切り換えセクタ1112に、通常／診断バケット送出切り換え信号1163を通常／診断バケットセクタ1113に、折り返しバケット送出切り換え信号1164を折り返しバケットセクタ1151に出力し、送信回路内のバケット送信経路を制御する。

【0027】送信バッファ1102は、プロセッサ内部バス13からバスデータ入力信号110を介して、OSなどがメモリ12上のバケット送信領域に作成したバケットを受け取るバッファである。同時に本バッファが使用中であるか否かを送信回路ステータス信号141を介してネットワークステータスレジスタ(NSR)103に対して報告を行う。

【0028】診断バケット作成回路1103は、送信回路制御回路1101から診断バケット作成信号1161を介して診断バケット作成命令を受けると、自動的に診断バケットを作成し、バケットをバケット送信経路に出力する。

【0029】バケットコード解析回路1104はバスデータ入力信号110から入力されたバケットのバケットコードのみを読み取って、内部バス13から入力されたバケットの種類が通常のバケットか診断バケットかを判定し、通常／診断切り換えスイッチ1111に送信バケットコード解析報告信号1171を出力し、バケットの送信経路を切り換える。

【0030】図4は図3と同様に、図2で述べたネットワークインターフェース回路14を構成する回路の一つである、受信回路102の内部構成を示す。図4において、受信回路102はネットワーク4からバケットを取り込む受信バッファ1201、受信バッファ1201が入力したバケットのフラグを読み取って、バケットの種類が通常バケットか診断バケットか折り返し診断バケットかを判定するバケットコード解析回路1202、受信バッファ1201やバケットコード解析回路1202など受信回路を構成する回路のステータスを集約し、ネットワークステータスレジスタ(NSR)103に報告する受信回路ステータス報告回路1203、バケットコード解析回路1202の結果により受信経路を切り換える折り返しバケット切り換えスイッチ1211及び通常／診断バケット切り換えスイッチ1212などを備える。

【0031】受信バッファ1201はネットワーク4からバケット受信信号121を介しバケットを取り込み、そのバケットを受信バケット経路(1)1221へ出力し、同時に受信バッファ1201内のバケット受信状況等の報告を受信バッファ状態報告信号1230を介して受信回路ステータス報告回路1203へ出力する。

【0032】パケットコード解析回路1202は、ネットワーク4からパケット受信信号121を介して入力された受信パケットのパケットコードのみを読み取って、その値により受信したパケットが通常パケットか受信診断パケットか送信診断パケットかの判定を行い、その結果を通常／診断パケット受信報告信号1240及び折り返しパケット受信報告信号1250を介して受信回路ステータス報告回路1203へ出力する。また同時に折り返しパケット切り換えスイッチ1211と通常／診断パケット切り換えスイッチ1212の制御を行い受信回路内部のパケット受信経路を切り換える。

【0033】受信回路ステータス報告回路1203は受信バッファ1201からの受信バッファ状態報告信号1230とパケットコード解析回路1202からの通常／診断パケット受信報告信号1240及び折り返しパケット受信報告信号1250を受け、その内容を受信回路ステータス信号142を介しネットワークステータスレジスタ(NSR)103へ出力する。

【0034】図5は本発明に係る並列計算機のプロセッサ間データ転送に使用するパケットのフォーマットを示す。図5においてパケットは、パケットの始まりを示すパケットヘッダーと内容を示すパケットデータから成る。パケットヘッダーはパケットの種類を示すパケットコード、パケットの大きさを示すパケットレングス、送信元を示す送信元プロセッサアドレス及び受信先を示す受信先プロセッサアドレスで構成される。

【0035】パケットコードは、前に述べたようにパケットの種類を示すが、具体的には通常の演算時に使用する通常パケットと自プロセッサが他のプロセッサに対して診断をする送信診断パケットと他のプロセッサからの診断パケットの返事である受信診断パケットによって値が変えられ、本実施例では通常通信パケットは“00”、送信診断パケットは“01”、受信診断パケットは“10”とする。

【0036】パケットレングスは、パケットに含まれるパケットデータの数を示し、診断パケットの場合解析を容易にする為、パケットデータを2個に設定する。

【0037】送受信プロセッサのアドレスは、並列計算機立ち上げ時に各プロセッサ毎に違った値がサービスプロセッサSVPより割り当てられているので、その値を参照して設定する。

【0038】パケットデータは、診断用パケットの場合ダミーデータとし、解析を容易にする為、第一パケットデータをオール“0”に、第二パケットデータをオール“F”に設定する。

【0039】図6は本発明に係る並列計算機であり、n個の演算プロセッサと診断専用のプロセッサをネットワークに接続し、プロセッサ間データ転送とネットワークを診断する診断パケットの送受信を行う並列計算機の構成を示す。図6において、1、2は演算プロセッサ(P

U(1), ..., PU(n))を示し3は診断専用プロセッサ(PU(n+1))を示す。各プロセッサ1, 2, 3はCPU11, 21, 31とメモリ12, 22, 32及びネットワークインターフェース回路14, 24, 34を備え、それぞれがプロセッサ内部バス13, 23, 33を介して接続される。各プロセッサ1, 2, 3は、データ転送信号10, 20, 30によりプロセッサ間結合のネットワーク4に接続される。5はサービスプロセッサ(SVP)であり、各プロセッサ1, 2, 3にプロセッサ制御信号6を介して接続される。

【0040】続いて本実施例に係る並列計算機の動作を述べる。

【0041】システム立ち上げ時、サービスプロセッサ(SVP)5が本並列計算機の構成情報を元に、プロセッサ制御信号6を介して各プロセッサ1, 2に並列計算機を構成するネットワーク4上でのアドレスやネットワーク構成情報などを配布し配布終了後、各プロセッサ内部が立ち上げ処理を開始する。各プロセッサ1, 2は装置立ち上げ処理に入ると、ネットワークインターフェース回路14に対してセルフテストの要求を発行する。その為に、ネットワークインターフェース回路14内部のネットワークコントロールレジスタ(NCR)に、自プロセッサに対して送信診断パケットを送信する命令を発行する。その命令を受けるとネットワークインターフェース回路14内部の送信回路101が自プロセッサ1宛の送信診断パケット自動診断パケット作成回路1104にて作成し、ネットワーク4に送信し同時に、この送信診断パケットは診断回路106に格納しておく。ネットワーク4に送信後、送信診断パケットの受信先プロセッサアドレスが自プロセッサ1宛になっている為、送信診断パケットがネットワーク4からそのままネットワークインターフェース回路14内に受信される。パケットをネットワークインターフェース回路14内の受信回路102が受信すると、そのパケットは送信診断パケットの為、送信して来たプロセッサ側のアドレスへ送り返す処理を行う。受信回路102内のパケットコード解析回路1202が送信診断パケットと解析すると折り返しパケット切り換えスイッチ1211を折り返しパケット処理側に切り換え、パケットヘッダー書換回路107でパケットコードを受信診断パケットのコード“10”に書き換えた後、送信回路101へパケットを転送し、送信回路101から再びネットワーク4へと受信診断パケットを送信する。ネットワーク2がこのパケットを受けると、送信先プロセッサアドレスが同じ為、再び自プロセッサ1に対して、本パケットを送り返す。自プロセッサ1が、受信診断パケットをネットワークインターフェース回路14内部の受信回路102に受信すると、パケットコード解析回路1202がパケットコードを解析し、受信診断パケットと判断する。すると、折り返しパケット切り換えスイッチ1211と通常／診断パケット切り換

えスイッチ1212を制御し、受信したパケットをネットワーク診断回路106へ転送する。ネットワーク診断回路106が受信診断パケットを受けると、先程格納しておいた送信診断パケットと比較し、パケットが正しければネットワーク及び自プロセッサ1のネットワークインターフェース回路14が正常であることをネットワークステータスレジスタ(NSR)103に対して報告する。これらの動作が終了するとCPU11がネットワークステータスレジスタ(NSR)を参照し、問題が無ければセルフテストを正常終了するが、送受信診断パケットが受信されなかった場合にはタイムアウトを返し、無事受信されても診断結果が悪ければ、診断パケットコンペアーエラーを返す。セルフテストでこれらの不良が発生すると、サービスプロセッサ(SVP)5に対して自プロセッサ1が不良である事を報告する。

【0042】セルフテストが正常終了すると自プロセッサ1から他のプロセッサ2に対しての診断を行う。これは、パケットの作成方法はセルフパケットと同様に診断パケット作成回路1104によってパケットを作成し、今度はセルフパケットのように受信先アドレスを自プロセッサ1宛ではなく、他のプロセッサ2宛のアドレスに設定し、ネットワークインターフェース回路14内部の送信回路101がネットワーク4を介して他のプロセッサ2に送信する。他のプロセッサ2が自プロセッサ1からの送信診断パケットをネットワークインターフェース回路24に受信すると、セルフテストと同様にパケットコードを受信診断パケットに変換し、且つ送信元アドレスと受信先アドレスのスイッチを行って、ネットワーク4を介し自プロセッサ1に送り返す。そして自プロセッサ1が受信診断パケットを受信すると、ネットワークインターフェース回路14内部の受信回路102は受信診断パケットを、ネットワークインターフェース回路14内部のネットワーク診断回路106に転送し、先ほど送信した送信診断パケットと比較し、比較結果をネットワークステータスレジスタ(NSR)103に報告する。その後、CPU11がネットワークステータスレジスタ(NSR)103の値を読み取り、今回診断したプロセッサ2の正常/異常を判定する。他のプロセッサ2へ診断が終了すると、次のプロセッサへの診断パケット送信を行い、これを全プロセッサに対して順番に繰り返す。

【0043】各プロセッサ1, 2がそれぞれ全プロセッサに対する診断を終了すると、システムの稼働に入り、OS等のブートを開始する。

【0044】システム稼働時の診断は、従来の技術であるユーザーコマンドやOS等からのソフトウェアルーチンによる診断パケットの送受信をするものと、ネットワークインターフェース回路14が自動的に診断パケットを作成しネットワークへ送受信するものと2種類がある。従来の技術であるユーザーコマンドからのソフトウェアルーチンによる診断パケットの送受信は、まずメモ

リ12上の送信パケット領域にダミーパケットを作成し、診断を行いたいプロセッサに対して通常パケットとして送受信を行い、送信したパケットと受信したパケットをOSがソフトウェア的に解析/判断する方法と、送信診断パケットとしてメモリ上にパケットを作成し、他のプロセッサ2へ送信し、他のプロセッサ2が受信診断パケットに変換して折り返しパケット送信したものを自プロセッサ1が受信してネットワークインターフェース回路14がハードウェア的に解析し、その結果をOSが参照して判断する方法とがある。

【0045】ネットワークインターフェース回路14が自動的に診断パケットを作成しネットワークへ送受信する方法には、OSがネットワークインターフェース回路14に対し、診断パケットを自動生成して診断を行う命令を発行し、ネットワークインターフェース回路14内部の診断パケット作成回路1103が診断パケットを自動的に生成し、他のプロセッサ2に対して診断パケットの送受信を行い、ネットワーク診断回路106でハードウェア的に解析し、その結果をOSが参照し判断する方法と、OSが前もってネットワークインターフェース回路14を、ネットワークが未使用であれば診断パケットを自動生成して診断を行うモードに設定し、自プロセッサ1のネットワークインターフェース回路14に対してCPU11がアクセスしていない時、診断パケットをネットワークインターフェース回路14内部で診断パケットを自動生成し、他のプロセッサ2に対して診断パケットの送受信を行い、ネットワーク診断回路106でハードウェア的に解析し、その結果が異常であればOSに対して報告する方法と、OSがネットワークインターフェース回路14を、一定期間毎に診断パケットを発行するモードに設定し、ネットワークインターフェース回路14内部でタイマー監視回路105のタイマー信号150を参照して一定期間毎に診断パケットを自動生成し、他のプロセッサ2に対して診断パケットの送受信を行いネットワーク診断回路106でハードウェア的に解析し、その結果が異常であればOSに対して報告する方法がある。

【0046】続いて、今までに述べた並列計算機のシステムに追加して、通常の演算を行う複数のプロセッサ1, 2以外に診断専用プロセッサ3をネットワーク4に接続した場合の動作を以下に述べる。システム立ち上げ時、サービスプロセッサ(SVP)5が本並列計算機の構成情報を元に、プロセッサ制御信号6を介して各演算プロセッサ1, 2と診断専用プロセッサ3に並列計算機を構成するネットワーク4上でのアドレスやネットワーク構成情報などを配布し配布終了後、各プロセッサ内部での処理を開始する。各演算プロセッサ1, 2と診断専用プロセッサ3は装置立ち上げ処理に入ると、各プロセッサ内部のネットワークインターフェース回路に対してセルフテストの要求を発行し、前に述べたものと同様セ

セルフテスト処理を実行する。セルフテスト処理が終了すると、演算プロセッサ1, 2はサービスプロセッサ(SVP)5からのシステム立ち上げ処理終了の報告が入るまでしばらく待ち状態になる。ここで診断専用プロセッサ3のみが動作し、各演算プロセッサ1, 2に対する診断処理を行う。診断専用プロセッサ3が全プロセッサに対する診断処理を終了すると、診断専用プロセッサ3がサービスプロセッサ(SVP)5に全プロセッサに対する診断処理終了の報告を行い、サービスプロセッサ(SVP)5がそれを受けるとサービスプロセッサ(SVP)5は全プロセッサにシステム立ち上げ処理終了の報告を行う。全プロセッサがシステム立ち上げ処理終了の報告をサービスプロセッサ(SVP)5から受けると全プロセッサはOS等のブート処理に入りシステム稼働開始となる。本構成で、システム運用中のネットワーク診断テストは、演算用プロセッサ1, 2からの診断パケット送受信は行わず、ネットワーク診断専用プロセッサ3のみが各演算用プロセッサ1, 2に対する診断パケットの送受信とセルフテストを行う。このネットワーク診断専用プロセッサ3によるネットワークの診断処理において他のプロセッサの故障等が発見されると、演算プロセッサ1, 2は診断処理を行っていないので、自分からは故障プロセッサの存在を把握することができないが、診断専用プロセッサ3がネットワークの診断処理を行うことによって診断専用プロセッサ3は故障プロセッサの発見が可能である。診断専用プロセッサ3が演算用プロセッサの故障を発見すると、ネットワーク4を介して全演算プロセッサ1, 2に対して、通常のパケットを使用して、ブロードキャスト転送を行い、報告する。この診断専用プロセッサは常に各自プロセッサ1, 2及び3に対する診断パケットの送受信を行っており、通常のパケットは診断パケット解析結果で故障プロセッサが発見されたときの各演算用プロセッサ1, 2に対するブロードキャスト報告の時のみ行う。

【0047】

【発明の効果】以上に述べたように、本発明によれば、複数のプロセッサを接続したネットワークによりプロセッサ間データ転送を行う並列計算機においてシステム立ち上げ時またはシステム運行中、ネットワークや他のプロセッサに対してネットワークを構成するハードウェアの故障等を診断する診断専用のパケットを送受信することにより障害の早期検出が可能になる為、計算機システムの稼働率が上がる。又、システムの運行中、通常データ転送路と切り換えて診断パケットの送受信を行う為、通常データ通信と衝突することなく安全であり、又ネットワーク未使用時や定期的な診断パケットの送受信を行う為、ネットワークの使用効率が向上する。

【0048】さらに演算プロセッサ以外に診断専用のプロセッサを設けることにより、頻繁な診断パケット送受信によるCPUやシステム全体の処理能力の低下を防ぐ

事ができる。

【図面の簡単な説明】

【図1】本発明の実施例に係る並列計算機の構成の概要を示す。

【図2】本発明の実施例に係る並列計算機のネットワークインターフェース回路の構成を示す。

【図3】本発明の実施例に係るネットワークインターフェース回路内部の送信回路の構成を示す。

【図4】本発明の実施例に係るネットワークインターフェース回路内部の受信回路の構成を示す。

【図5】本発明の実施例に係る並列計算機のプロセッサ間データ転送に用いるパケットのフォーマットを示す。

【図6】本発明の実施例に係る並列計算機の構成の概要を示す。

【符号の説明】

1…プロセッサ(PU(1))、2…プロセッサ(PU(n))、3…プロセッサ(PU(n+1))、4…ネットワーク、5…サービスプロセッサ(SVP)、6…プロセッサ制御信号、7…ネットワーク制御信号、10…データ転送信号(PU(1))、11…CPU(PU(1))、12…メモリ(PU(1))、13…プロセッサ内部バス(PU(1))、14…ネットワークインターフェース回路(PU(1))、20…データ転送信号(PU(n))、21…CPU(PU(n))、22…メモリ(PU(n))、23…プロセッサ内部バス(PU(n))、24…ネットワークインターフェース回路(PU(n))、30…データ転送信号(PU(n+1))、31…CPU(PU(n+1))、32…メモリ(PU(n+1))、33…プロセッサ内部バス(PU(n+1))、34…ネットワークインターフェース回路(PU(n+1))、40…ネットワーク回路障害報告ラッチ、101…送信回路(PU(1))、102…受信回路(PU(1))、201…送信回路(PU(n))、202…受信回路(PU(n))、301…送信回路(PU(n+1))、302…受信回路(PU(n+1))、103…ネットワークステータスレジスタ(NSR)、104…ネットワークコントロールレジスタ(NCR)、105…タイマー監視回路、106…パケット診断回路、107…ヘッダー書換回路、110…バスデータ入力信号、111…パケット送信信号、112…送信診断パケット出力信号、113…折り返しパケット入力信号、120…バスデータ出力信号、121…パケット受信信号、122…受信診断パケット出力信号、123…折り返しパケット出力信号、130…ネットワークコントロール信号、131…送信回路制御信号、140…ネットワークステータス信号、141…送信回路ステータス信号、142…受信回路ステータス信号、143…診断回路ステータス信号、150…タイマー信号、1101…送信回路制御回路、1102…送信バッファ、1103…パケットコード解析回路、1104…診断パケット作成

回路、1111…通常／診断切り換えスイッチ、1112…診断
 パケット作成手段切り換えセレクタ、1113…通常／診断
 パケットセレクタ、1114…通常／折り返しパケットセレ
 クタ、1121…送信パケット転送路(1)、1131…通常パケ
 ット転送路、1132…診断パケット転送路、
 1133…自動作成診断パケット転送路、1141…送信パ
 ケット転送路(2)、1151…診断パケット作成信号、1152
 …診断パケット作成手段切り換え信号、1153…通常／診
 断パケット送出切り換え信号、1154…折り返しパケット

切り換え信号、1161…通常／診断パケット入力切り換え
 信号、1201…受信バッファ、1202…
 パケットコード解析回路、1203…受信回路ステータス報
 告回路、1211…折り返しパケット切り換えスイッチ、12
 12…通常／診断パケット切り換えスイッチ、1221…受信
 パケット経路(1)、1222…受信パケット経路
 (2)、1230…受信バッファ状態報告信号、1240…通常
 ／診断パケット受信報告信号、1250…折り返しパケット
 受信報告信号。

【図1】

図 1

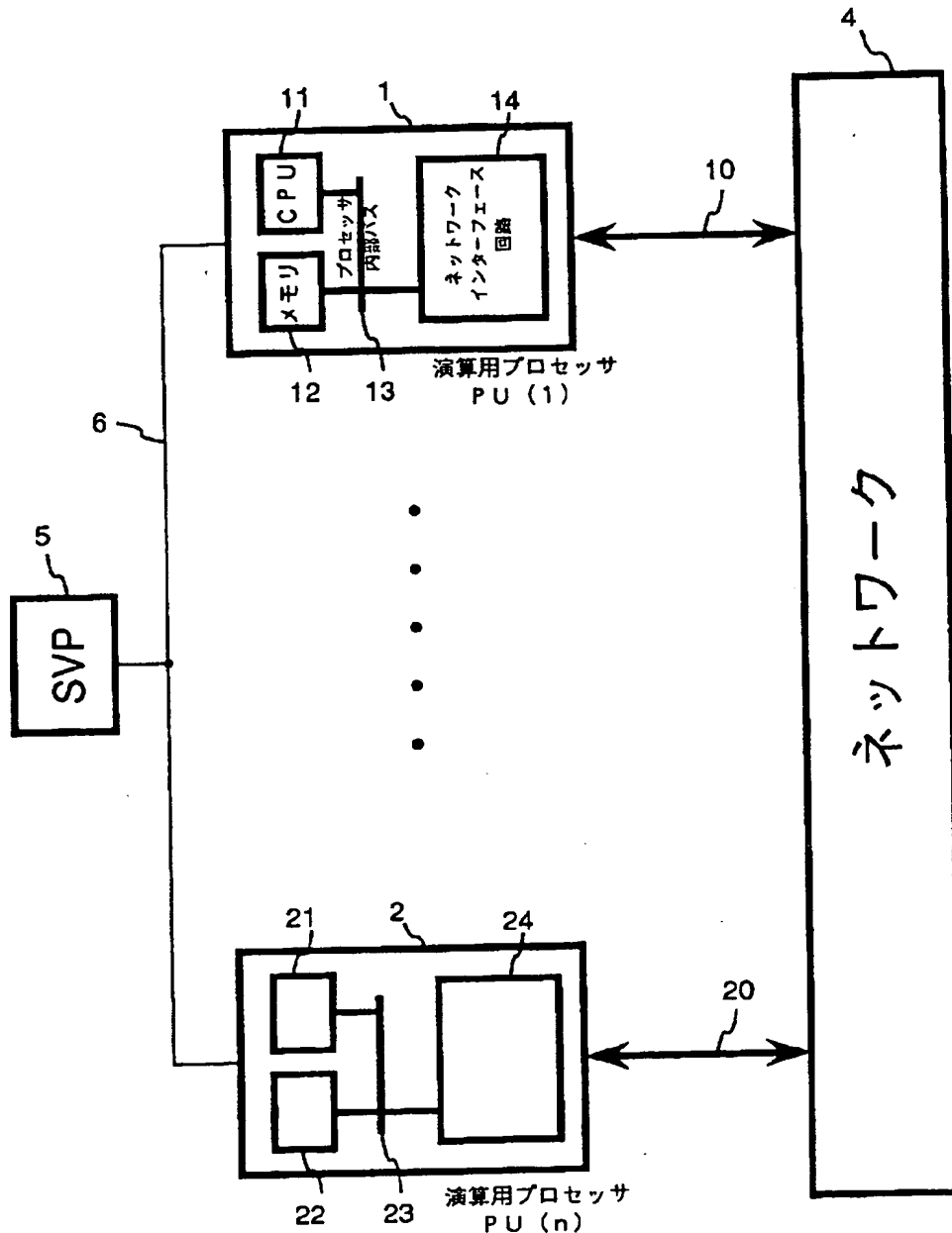
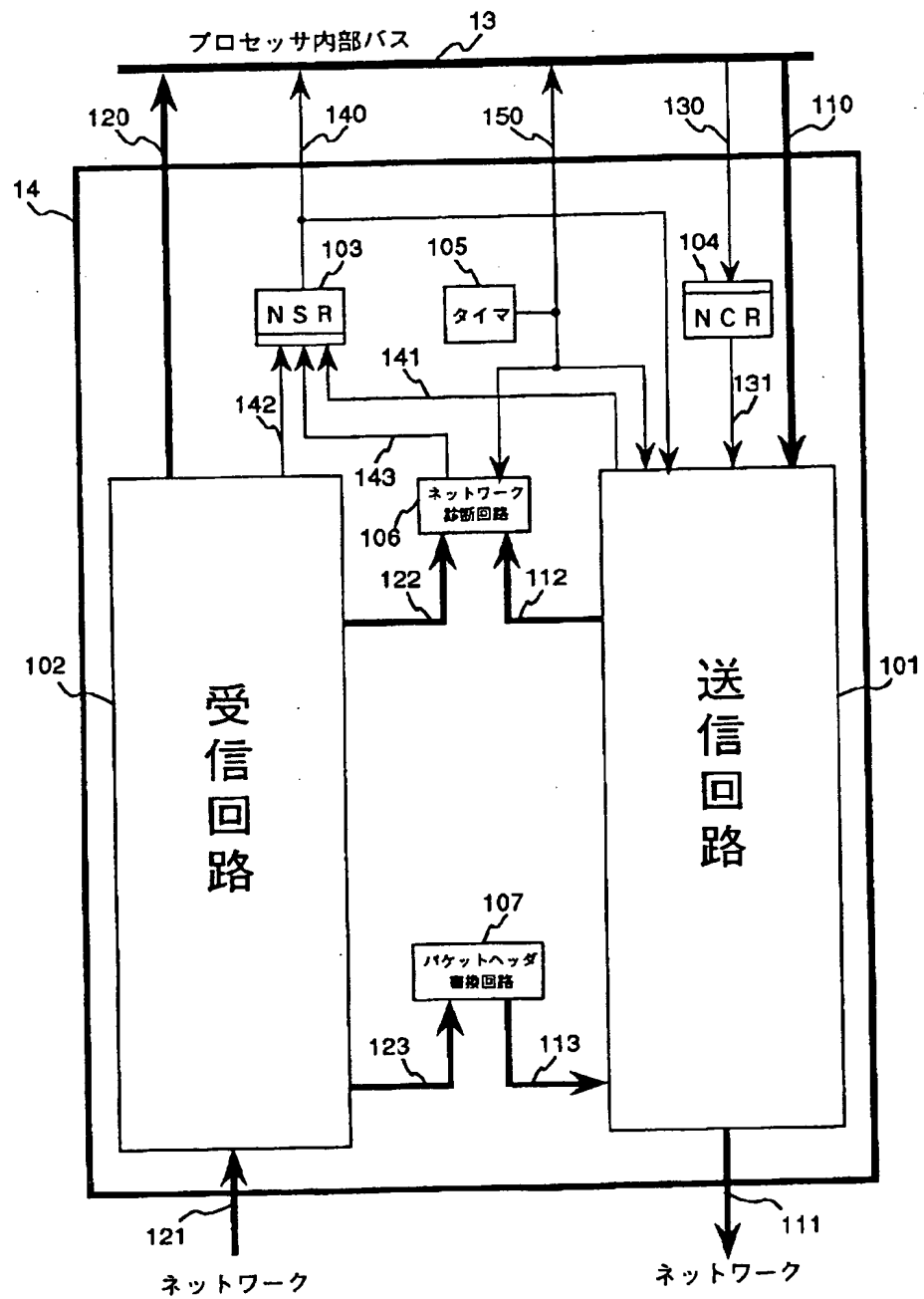
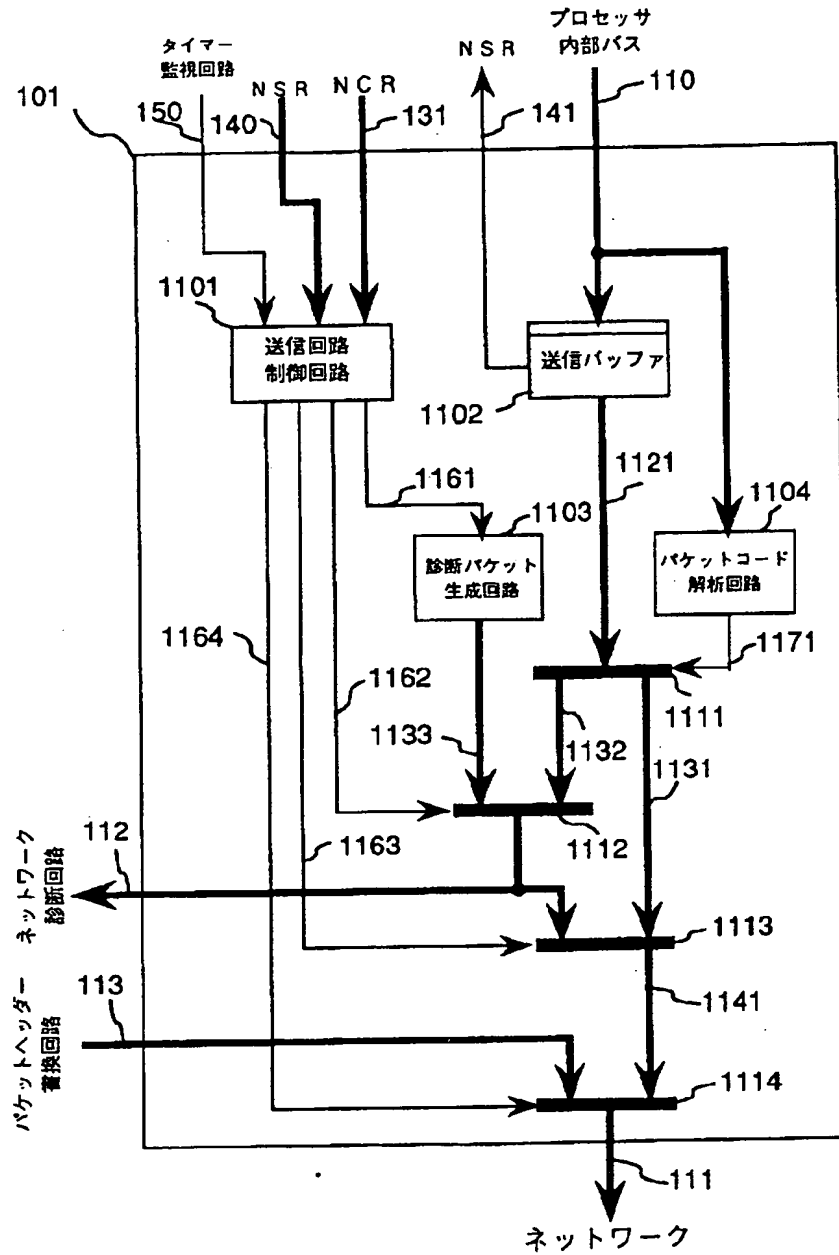


图 2



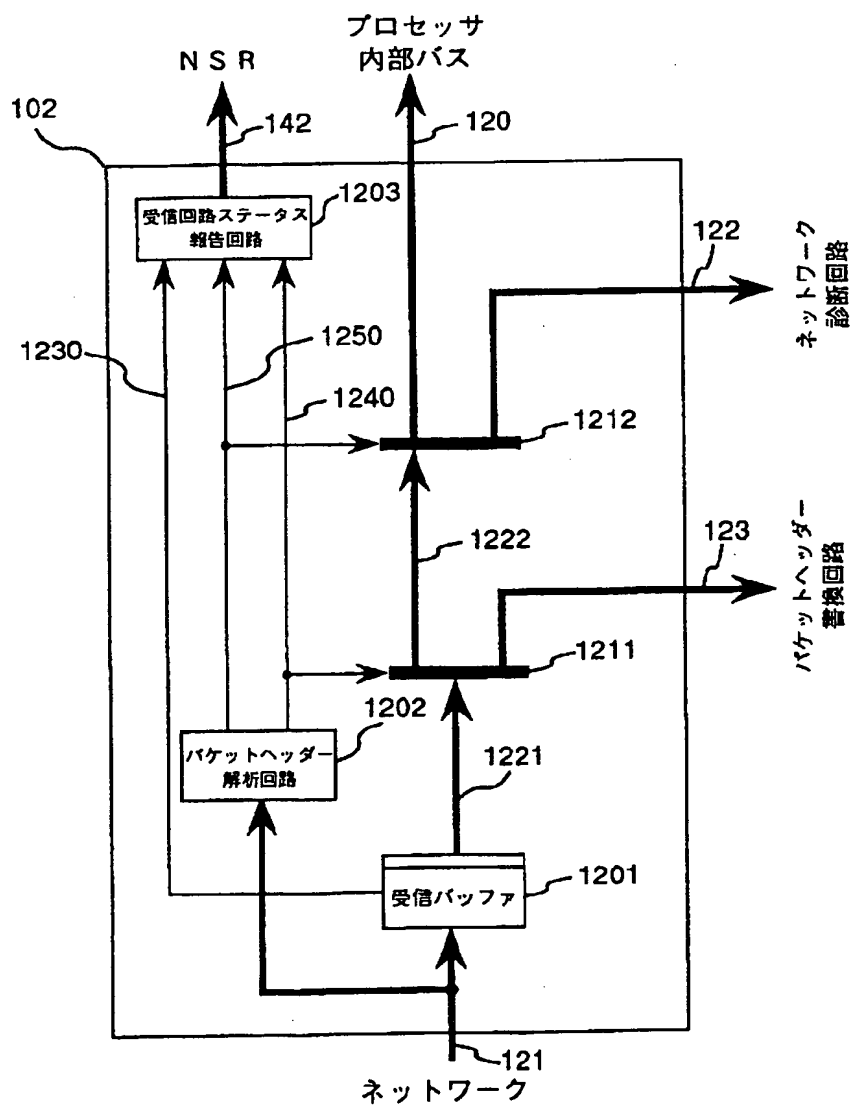
【図3】

図-3



【図4】

図 4



【図5】

図 5

パケット コード	パケットレングス	送信元プロセッサアドレス	受信先プロセッサアドレス
パケットデータ (0)			
パケットデータ (1)			
パケットデータ (2)			
・ ・ ・			
パケットデータ (n)			

【図6】

図 6

